

Abstract of the Disclosure

A method and apparatus for performing discriminative training of, for example, call routing training data (or, alternatively, other classification training data) which improves the subsequent classification of a user's natural language based requests. An initial scoring matrix is generated based on the training data and then the scoring matrix is adjusted so as to improve the discrimination between competing classes (*e.g.*, destinations). In accordance with one illustrative embodiment of the present invention, a Generalized Probabilistic Descent (GPD) algorithm may be advantageously employed to provide the improved discrimination. More specifically, the present invention provides a method and apparatus comprising steps or means for generating an initial scoring matrix comprising a numerical value for each of a set of n classes in association with each of a set of m features, the initial scoring matrix based on a set of training data and, for each element of said set of training data, based on a subset of said features which are comprised in the natural language text of said element of said set of training data and on one of said classes which has been identified therefor; and based on the initial scoring matrix and the set of training data, generating a discriminatively trained scoring matrix for use by said classification system by adjusting one or more of said numerical values such that a greater degree of discrimination exists between competing ones of said classes when said classification requests are performed, thereby resulting in a reduced classification error rate.